

REMARKS

Enclosed herewith is a Substitute Specification in which the specification as filed has been amended in various places to correct typographical and grammatical errors, and also to add section headings.

In support of the above, enclosed herewith is a copy of the specification as filed marked up with the above changes.

The undersigned attorney asserts that no new matter has been incorporated into the Substitute Specification.

The claims have been amended to more clearly define the invention as disclosed in the written description. In particular, the claims have been amended for clarity.

Claim 3 has further been amended to clarify the expression "a text color is fewer present than a background color." In particular, claim 3 now recites "a number of picture units of the text color is fewer than a number of picture units of the background color." This finds support in the Substitute Specification on page 16, line 19 to page 17, line 17, (paragraphs [0045] - [0047]).

Applicants believe that the above changes answer the Examiner's 35 U.S.C. 112, paragraph 2, rejection of claim 3, and respectfully request withdrawal thereof.

The Examiner has rejected claims 1, 7 and 8 under 35 U.S.C. 102(e) as being anticipated by U.S. Patent 6,061,035 to

Suzuki et al. The Examiner has further rejected claim 2 under 35 U.S.C. 103(a) as being unpatentable over Suzuki et al. in view of U.S. Patent 5,956,468 to Ancin. Moreover, the Examiner has rejected claim 3 under 35 U.S.C. 103(a) as being unpatentable over Suzuki et al. in view of U.S. Patent 6,038,340 to Ancin et al. and further in view of Ancin. In addition, the Examiner has rejected claim 4 under 35 U.S.C. 103(a) as being unpatentable over Suzuki et al. in view of U.S. Patent 6,148,102 to Stolin. Furthermore, the Examiner has rejected claim 5 under 35 U.S.C. 103(a) as being unpatentable over Suzuki et al. in view of the article "Thresholding and Enhancement of Text Images for Character Recognition" by W.W. Cindy Jiang, and further in view of U.S. Patent 5,781,658 to O'Gorman. Finally, the Examiner has rejected claim 6 under 35 U.S.C. 103(a) as being unpatentable over Suzuki et al. in view of the Jiang article.

The Suzuki et al. patent discloses a display control apparatus having a display of text data and image data and a display control method for the same, in which in a signal it is detected whether a region is or is not a text region. If the region is a text region, scaling is performed with a scale down percentage of 66%, and if the region is not a text region, scaling is performed with a scale down percentage of 50%.

The subject invention relates to a process for improving text in a signal. To that end, the signal is analyzed to detect text regions (as in Suzuki et al.). However, in the subject

invention, the entire signal is then subjected to scaling in order to match the signal with the display. Then, the scaled signal is post processed in dependence on the detected text regions.

Applicants submit that Suzuki et al. does not show or suggest these steps of the inventive method, i.e., scaling the entire signal after detection of the text regions, and then post-processing the scaled signal in dependence on the previously detected text regions.

The Ancin patent discloses a document segmentation system in which a signal having both text and image regions is analyzed. The text regions are detected (i.e., "The method and apparatus then detect dark text pixels on a light background region of a document and assign segmentation labels to each pixel.") As further stated in the Abstract, "The pixel labels are post-processed using a plurality of syntactic rules to correct mislabeling of pixels." Further, "Pixels identified as being in the background region of the document are assigned a white label and pixels identified as being in the text region are assigned a black label."

Applicants submit that Ancin only changes the text to black and the background to white. Ancin neither shows nor suggests that this should be done in a pre-scaling step, and that in a post-scaling step, the text and background should be returned to their previous colors.

Furthermore, Applicants submit that Ancin does not supply that which is missing from Suzuki et al., i.e., scaling the entire signal after detection of the text regions, and then post-processing the scaled signal in dependence on the previously detected text regions.

The Ancin et al. patent discloses a system and method for detecting the black and white points of a color image, in which, in a particular region, the number of black pixels is compared to the number of white pixels to determine whether that region is a text region. However, Applicants submit that Ancin et al. or Ancin, either individually or collectively, do not supply that which is missing from Suzuki et al., i.e., scaling the entire signal after detection of the text regions, and then post-processing the scaled signal in dependence on the previously detected text regions.

The Stolin patent discloses recognizing text in a multicolor image, in which, in detecting text, "the image is separated into multiple blocks. Color distribution of each of the blocks are analyzed, and blocks having two main colors are identified. The two-color blocks having similar colors are grouped into two-color zones, and text in the two-color zones are identified." However, Applicants submit that Stolin does not supply that which is missing from Suzuki et al., i.e., scaling the entire signal after detection of the text regions, and then post-

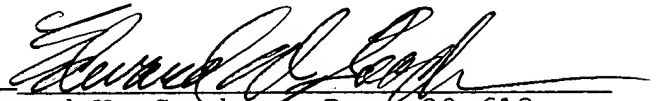
processing the scaled signal in dependence on the previously detected text regions.

The Jiang article discloses "Each sub-pixel in the interpolated graytone image has an associated threshold ... If a sub-pixel intensity is larger than its threshold, binarize it into a black sub-pixel; otherwise make it a white sub-pixel." However, as Applicants noted above with respect to the other references, Jiang does not supply that which is missing from Suzuki et al., i.e., scaling the entire signal after detection of the text regions, and then post-processing the scaled signal in dependence on the previously detected text regions.

In view of the above, Applicants believe that the subject invention, as claimed, is neither anticipated nor rendered obvious by the prior art, either individually or collectively, and as such, is patentable thereover.

Applicants believe that this application, containing claims 1-8, is now in condition for allowance and such action is respectfully requested.

Respectfully submitted,

by 
Edward W. Goodman, Reg. 28,613
Attorney
Tel.: 914-333-9611

CERTIFICATE OF MAILING

It is hereby certified that this correspondence is being deposited with the United States Postal Service as First-class mail in an envelope addressed to:

COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, VA 22313-1450

On June 16, 2003

By Burnett James



TEXT IMPROVEMENT

RECEIVED

JUN 20 2003

Technology Center 2600

BACKGROUND OF THE INVENTIONField Of The Invention

[0001] The invention relates to a method and device for text improvement.

5

Description Of The Related Art

[0002] The article "Thresholding and enhancement of text images for character recognition", by W.W. Cindy Jiang, IEEE, Proceedings of the international conference on acoustics, speech, and signal processing (ICASSP), NY, vol. 20, 1995, pp. 2395-2398, discloses a scheme which converts graytone text images of low spatial resolution to bi-level images of higher spatial resolution for character recognition. A variable thresholding technique and morphological filtering are used. It is stated that most optical character recognition systems perform binarization of inputs before attempting recognition, and that text images are usually supposed to be binary.

[0003] The article "A segmentation method for composite text/graphics (halftone and continuous tone photographs) documents", by S. Ochuchi et al., Systems and Computers in Japan, Vol. 24, No. 2, 1993, pp. 35-44, discloses that when processing composite documents for digital copy machines and facsimile which contain a mixture of text, halftone and continuous tone

photographs, ideally, the text portion can be separated from the graphics portion and more efficiently represented than the multi-bit pixel bitmap graphics representation.

[0004] Nowadays, digital display devices are more and more frequently matrix devices, e.g., Liquid Crystal Displays, where each pixel is mapped on a location of the screen having a one-to-one relationship between raster data and display's points. This technology implies the usage of a scaling system to change the format of the input video/graphic signal so that it satisfies the size of the device, i.e., the number of its pixels. The scaling block is based on a filter bank that performs pixel interpolation when the zooming factor is varying. Actually available solutions on the market apply an undifferentiated processing on the graphic raster that leads to results with unavoidable artifacts. Usually, low-pass filters reduce pixellation, also know as the seesaw effect, on diagonals, and prevent the signal ~~to suffer from~~ suffering from aliasing due to the sub-sampling, but they also introduce other annoying effects, such as, blurring the images. It depends on the content of the displayed signal, the relevance of the perceived artifacts and the kind of artifacts that have to be preferred as unavoidable.

SUMMARY OF THE INVENTION

[0005] It is, inter alia, an object of the invention to provide a simple text improvement for use with displays that require a

scaling operation. ~~To this end the invention provides a text improvement as defined in the independent claims. Advantages embodiments are defined in the dependent claims.~~

[0006] Starting from the above-mentioned observations, a novel
5 technique is provided here that is able to take into account the
image content and to apply an ad hoc post-processing only where it
is required. ~~So~~Hence, in accordance with the present invention,
text improvement after the scaling operation is based on text
detection before the scaling operation. The processing is active
10 only in presence of text region. A viable area of application of
this invention is the text readability improvement in the case of
LCD devices, when, and it is usually the case, we do not want to
affect other parts of the displayed signal.

[0007] A remarkable characteristic of the technique presented
15 here is its really low computational complexity. This aspect
determines a high effectiveness in terms of cost/performances
ratio. In fact, the insertion of the proposed algorithm into the
other circuitry that carries out all the digital processing needed
for ~~resize~~ resizing the matrix display device input, presumably
20 ~~raises~~ raises the display quality, according with the average user
perception, without ~~affecting~~ considerable affecting its cost.

[0008] It is noted that while in one embodiment, a binarization
takes place, this binarization is only carried out in regions where
text has been detected, while in the prior art, the binarization is

a preliminary step to be carried out before characters can be recognized.

[0009] These and other aspect of the invention will be apparent from and elucidated with reference to the embodiments described hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] In the drawings:

[0011] Figs. 1-3 illustrate the operation of a morphological filter; and

[0012] Fig. 4 shows a block diagram of a system in accordance with the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0013] The invention proposes the design of a text detection algorithm, together with a post-processing block, for text enhancement. It will be shown that the invention significantly improves the performance in terms of content readability and leads to good perceptual results of the whole displayed signal, while keeping really low the computational complexity of the total scaling system.

[0014] The organization of the remainder of this document is as follows. First the general scaling problem and the current available algorithms are briefly summarized. Thereafter concepts concerning format conversion by a non-integer factor will be

introduced. Successively, the post-processing block, characterized by the thresholding operation and the morphological filter, will be summarized and its features will be described. Finally the text search strategy will be presented and the detection algorithm and its cooperation with the previously introduced post-processing block will be elucidated.

The ~~general framework~~ General Framework

[0015] Resizing pictures into a different scale requires format conversion. This operation involves well-known re-sampling theory, and classical filter procedures are currently used to accomplish ~~with~~ it. Filters avoid aliasing problems in frequency, freeing room for the repetitions introduced by the sampling operation in the original domain. Among the interpolation filter families, polynomial interpolators of first order are commonly used, in which the reconstructed pixel is a weighted mean of the nearest pixels values. These kinds of filters are also called Finite Impulse Response (F.I.R.) filters.

[0016] Inside standard display devices, the format conversion problem is usually faced with linear filtering too. A particularly simple class of F.I.R. filters reconstructs pixels in between two available ones tacking the value on the line joining these two adjacent points.

[0017] There are many other possible techniques. For example, pixel repetition or polynomial interpolation with more complexes

weighting functions. The quality perception of images processed with these different solutions is generally not really high, there are impairments and artifacts that are not completely avoidable. This consideration implies that some compromise is due in order to reach an acceptable or, at ~~the~~ best, a satisfactory cost/performance ratio.

[0018] In the past, the simplest solution solved the problem using pixel repetition. A more recent solution, see the Philips scaler PS6721, still uses linear filtering but with a slight different shape of the impulse response, to improve the transition steepness.

[0019] Measuring the rise time of the step response is a classical way to assess the performance of the interpolator in presence of an edge. In fact, ~~low~~-low-pass filters affect edge steepness and a smooth steepness is perceived as a blurring effect.

[0020] Moreover, the actual impact of this annoying artifact depends on the kind of displayed signal. Actually, in case of natural images, a blurring effect could be tolerated in a certain measure.

[0021] Whereas, for artificial patterns, a slight smoothness effect is recommended only when the content requires approaching a natural impression (this is the case of virtual reality and 3D games). In this case, filtering is used as an anti-aliasing process. For the same reason, these kinds of filters are used on text/characters to avoid the pixellation effect, also know as the

seesaw effect, on diagonals. Interpolation filters are anti-aliasing filters too, because they reduce the highest frequency of the input signal. Moreover, supposing to have a black text on a white background, the amount of gray levels introduced by this kind of filters should be a less percentage of the black quote. If it is not the case, we have an artifact instead of a picture improvement, and images are perceived as blurred. For instance, when bilinear interpolation as well more sophisticated filters, like the bicubic ones, are used on small characters (the commonly used size 10÷12 points) and thin lines, they appear defocused. In all these cases, it seems better to use no filters at all, at least no ~~low~~-low-pass filters as are actually available.

[0022] Starting from the above consideration, we can conclude that, because format conversion requires resampling, so that the filtering process is unavoidable, to accomplish with the above issue we have to find out some other solution. In case of text, a simple idea is to apply a post-processing block after the scaler to clean all the gray levels where characters are detected. Because of the scale change, this operation could not be performed using only a simple threshold block. In fact, threshold is a non-linear operator that introduces ~~not~~-non-uniform patterns when it converts gray levels characters to binary values, another kind of artifact that is highly noticeable. Morphological filters are an interesting class of operators that are able to change ~~not~~-non-regular patterns

into more regular ones. They will be introduced in a following section.

Format ~~conversion by a rational factor~~Conversion By A Rational
5 Factor

[0023] In today's digital display devices, images are frequently represented with a matrix of pixels ~~so that~~requiring a fixed picture format ~~is required~~. When a signal with a different format arrives at the input of a matrix display, format conversion is
10 unavoidable. ~~Supposing~~If a graphic card had generated the signal, ~~than~~then selecting a different graphic format, instead of the one used by the display, depends on the requirement of the software application running. At the moment, it is not advisable to
15 constraint the graphic card output only with the requirement of the display.

[0024] We recall that standard today's graphic formats are VGA, SVGA, XGA, SXGA and higher. Format conversion between these raster sizes requires, in almost all ~~the~~ cases, a rescaling by a rational factor. This, by itself, leads to a sensible degradation of the
20 resampled picture. In fact when, for example, we need to change format from VGA at display's input to XGA at display's output, the factor involved would be $8/5$, equal to 1.6 times the size of the original picture. This format conversion ratio would clearly require a sub-pixel resolution, but with standard linear filtering

techniques, this is not ~~be allowed~~possible without paying a high blurring cost.

[0025] Let $s(i,j)$ be the input signal at position (i,j) in the input grid, and $\hat{s}(\tilde{i},\tilde{j})$ be the signal after format conversion at position (\tilde{i},\tilde{j}) in the thicker output grid. ~~Supposing to have~~In the case of a rescaling from VGA to XGA, i.e., by an 8/5 factor, for every 5 pixels at the input of the sampler there will be 8 pixels at its output. A rescaling by a rational factor conceptually relies on an intermediate "super-resolution" grid obtained using a zooming factor equal to the numerator, in the example 8. In this case, the "super-resolution" grid will be eight times thicker than the original one. Tacking two input values, $s(i,j)$ and $s(i+1,j)$, on the same line j at position i and $i+1$, the interpolated value $\hat{s}(\tilde{i},\tilde{j})$ will be positioned in between the two original values in the super-resolution grid, i.e., in one of the eight possible positions available on the grid. ~~We express this fact~~This fact is expressed by the following equation:

$$\hat{s}(i + \frac{k}{8}, j) = w_1 \cdot s(i, j) + w_2 \cdot s(i+1, j) \quad \forall k \in [0...7]$$

Where, for the linear interpolator:

$$\begin{cases} w_1 = \delta \\ w_2 = 1 - \delta \end{cases}$$

k is the position of the pixel in the dense grid, the position ~~is~~ also being called filter phase. In a linear filter, the position is $\delta \propto k$, and δ is the distance between the pixel to be interpolated with respect to the two adjacent original ones. The signal at the

5 output grid will be obtained tacking values on a grid 5 times weaker. The sub-sampled signal at the output is expressed as ~~it~~ follows:

$$s^*(i + 5 \cdot \frac{k}{8}, j) = w_1 \cdot \hat{s}(i + \frac{\bar{k}}{8}, j) \quad \bar{k} \in [0 \dots 7]$$

10

Because the output grid is not a multiple of the input grid, often original pixel values will be lost and they will be replaced with an average value according ~~with what we said~~ to the above. If the input pattern ~~would be~~ is a black and white text, its pixels will be

15 frequently replaced by a weighted average of their values, a gray level.

~~Text improvement via thresholding~~ Improvement Via Thresholding

[0026] A threshold operator placed at the output of the scaling

20 filter will recover a black and white pattern, or, more generally, a bicolor one, choosing the threshold nearest value according to the following relationship:

25

$$\begin{cases} s_g^*(i + \frac{8}{5} \cdot k, j) = l_k & \text{if } s^*(i + \frac{8}{5} \cdot k, j) < \vartheta \\ s_g^*(i + \frac{8}{5} \cdot k, j) = l_w & \text{if } s^*(i + \frac{8}{5} \cdot k, j) \geq \vartheta \end{cases}$$

Where l_k is the black level and l_w ~~is~~ the white one level.

[0027] ~~We could notice~~ It should be noted that, in case of black/white and bicolor patterns, the threshold function could be integrated in the filter operator, setting l_k and l_w in accordance with the actual filter phase. In this way, the threshold operation recovers original bicolor levels from the interpolated ones according with theirs new positions. In regions where the amount of gray levels introduced is too ~~height~~ high, this simple operator improves the sharp edge perception. Anyway, this is paid with the introduction of irregular patterns. In the next section, ~~we will see how a~~ solution to this problem could be solved is presented.

Morphological ~~filtering algorithms~~ Filtering Algorithms

[0028] The introduction of mathematical morphology to solve the problem of text deblurring is due to the fact that a morphological filter, working both as a detector and as a non-linear operator, is able to eliminate gray levels without destroying the character regularity. Moreover, in case of bicolor patterns, a morphological filter is able to recover a specified regularity where required.

[0029] In general, the detector, called structuring element, is a small matrix (usually 2x2 or 3x3) ~~it~~ It can recognize a particular pattern on the data, in our case, the rasterized image's pixels at the display output, and ~~to~~ substitute that pattern with a different set of requested values.

~~Supposing to use [0030]~~ When the morphological filter is used after the threshold block, on a bi-level pattern, the structuring element will work as a binary mask on the underlying data, performing a set of logical operations between the bit of the running matrix and the bit of the scanned data. An output equal ~~at~~ to 1 will signify that a specified pattern has been identified.

[0031] A particular operator belonging to the morphological filter family, also called "diagonal" filter, applies the following set of logical operations to the data:

$$Y = X_4 \cup (P_1 \cup P_2 \cup P_3 \cup P_4)$$

$$P_1 = (X_4^c \cap X_7 \cap X_6^c \cap X_3)$$

$$P_2 = (X_4^c \cap X_3 \cap X_0^c \cap X_1)$$

$$P_3 = (X_4^c \cap X_1 \cap X_2^c \cap X_5)$$

$$P_4 = (X_4^c \cap X_5 \cap X_8^c \cap X_7)$$

Here, $X_0 \dots X_8$ is the set of data currently analyzed by the

structuring element, ~~besides~~. In addition, in case of binary data,

\cup is the classic logical OR operator and \cap is the classic logical AND operator. The structuring element orders the data in its framework as shown in Fig. 1.

[0032] The output, y , after the set of logical operations ~~above~~ introduced above, replaces the previous value at the origin of the data matrix, X_4 in ~~the figure~~ Fig. 1. ~~One can notice~~ It should be noted that, if the result of $P_1 \cup P_2 \cup P_3 \cup P_4$ is 0, ~~than-then~~ X_4 remains unchanged, if, instead, ~~if the result if 1~~ than, then X_4 is always replaced by 1.

[0033] Looking carefully, it will be evident that the set P_1 , P_2 , P_3 , P_4 of logical operations corresponds to the detection of

the patterns shown in Fig. 2. Patterns in Fig. 2 are diagonal patterns of black and white pixels, in case of binary images.

~~According with~~ In accordance with the above relations, when one of these configurations is found, ~~than~~ then, in the origin of the detected region identified by a circle in ~~the figure~~ Fig. 2, a 0 value is substitute by a 1. This operation, in terms of pattern effect, fills holes in diagonal configurations.

[0034] ~~One could notice~~ It should be noted that the same operation could be done, instead of using logical operators, with a LUT addressed by the configuration of bits in the structuring element. ~~Supposing to order~~ When the cells of the element are ordered according ~~with the above figure~~ to Fig. 2, this configuration has the following address:

$$LUT_{address} = X_8 X_7 X_6 X_5 X_4 X_3 X_2 X_1 X_0$$

where each X_i is correspondingly equal at 1 or 0 according with the value in the i^{th} position of the matrix. To fill holes, the LUT at position $XXX10X01X$, $01X10XXXX$, $X10X01XXX$, $XXXX01X10$, will be set at 1, in all of the other positions, it will be set at 0. Here X means "don't care".

[0035] From a conceptual point of view, because of the holes filling function of the "diagonal" structuring element, the set of operations described above on a structuring element, are equivalent

~~at to~~ changing a diagonal patterns, anyhow oriented in a matrix, with a uniform block. This concept is clarified in Fig. 3.

~~Block diagram of system embodiment~~ Diagram Of System Embodiment

5 [0036] ~~The drawing in~~ Fig. 4 shows a block diagram of the total system in which the main concept of the architecture for the detector and the ~~post post~~-processing block are ~~sketched~~drawn. An input image InIm s is applied to a Search Window part SW and Text Detector part Det. The input image InIm s, possibly modified in
10 some region by the text detector part Det, is applied to a scaler Scal, if recognized as text by the text detector part Det, such as the commercially available scaler PS6721. The scaled image from the scaler Scal is applied to a post-processing part Post-proc that produces the output image OutIm s*.

15

Search Window and Text Detector

[0037] The search window and text detector is a key operator. In fact, it depends on ~~it~~ if the input signal will be binarized and further processed or simply filtered with the linear scaler.
20 According to what ~~it~~ was previously ~~said~~stated, detection is specifically designed to recognize text patterns. When the required constraint imposed at the detector are not satisfied, the signal does not eventually benefit ~~of this~~from further processing step. Detection is performed with a local sensor that recognizes the

amount of colors in a small region. So, in principle, it works as a search window that scan the raster image to discover text areas.

[0038] To design it, satisfying a low memory cost, a fixed vertical width was used, equal to 3 lines on the domain of the original signal. Instead, its horizontal depth ~~is varying~~ varies according to the image characteristics, and it is based on a simple growth criterion defined using some intuitive assumptions on the text properties. Currently assumptions on graphic text are as ~~it~~ follows:

[0039] 1. ——— A text area is a two-color region in which text is one color and the other color is the background.

[0040] 2. ——— In a text area, a text color is perceptually fewer present than a background color.

[0041] 3. ——— A text region has a reasonable horizontal extension.

[0042] These assumptions determine the constraints on the patterns the detector recognizes as text regions. As ~~we can be~~ seen, neither filtered text nor not-uniform background are recognized as text regions. This is a reasonable assumption because the threshold operator in these cases would introduce more artifacts than benefits. Furthermore, the ~~not-imbalanced~~ percentage of the two colors prevents the detector from identifying, as text, two color regions with potentially dangerous patterns. An example

is the chess pattern, quite recurrent, for example, in window folder background. Finally, the third condition prevents ~~to identify~~ identifying as text regions small bicolor fragments of the raster signal, that could ~~be~~ presumably be border or other small pieces of graphic objects.

[0043] The conditions introduced above are used to define some parameters to adjust the behavior of the detector such that it could ~~may~~ reach the best performances. ~~Let we c~~Consider now the above introduced input raster signal $s(r,c)$ at position (r,c) . The search window will be indicated with $q(r,c)$, with (r,c) being the coordinates of the block's origin that identify its reference pixel in the image; the relative coordinates, identifying a cell in the search window, are referred to the block origin and they will be noted as (i,j) . Furthermore, the detector height and width will be indicated with h and w . ~~Whereas~~ While w is a varying parameter, on the contrary, h is fixed, to satisfy line memory constraints, and its value is currently $h = \bar{h} = 3$.

[0044] Let ~~be~~ N_c be the number of colors detected in the search window. According ~~with to~~ the previously described block growing, the width w will increase following this search strategy:

$$\begin{cases} w(k+1) = w(k) + 1 & \text{if } N_c \leq 2 \\ w(k+1) = w(k) = w & \text{if } N_c > 2 \end{cases}$$

$N_c > 2$ is the exit condition from the growing search strategy. When the exit condition is verified, the system will return the final block width w .

[0045] Together with the block growing process, two color counters will be incremented at each new step k . It ~~could~~ should be ~~notice-noted~~ that a step k corresponds to the evaluation of a new input pixel in the horizontal direction. Calling γ_1 the number of pixels with color c_1 and γ_2 the number of pixels with color c_2 , the counters will be upgraded ~~according~~ in accordance with the corresponding block growing step in the following manner:

$$\begin{cases} \gamma_1(\tau+1) = \gamma_1(\tau) + 1 & \text{if } q(i+w(k+1), j+h) = c_1 \text{ for } h=1...3 \\ \gamma_1(\tau+1) = \gamma_1(\tau) & \text{otherwise} \end{cases}$$

and

$$\begin{cases} \gamma_2(\tau+1) = \gamma_2(\tau) + 1 & \text{if } q(i+w(k+1), j+h) = c_2 \text{ for } h=1...3 \\ \gamma_2(\tau+1) = \gamma_2(\tau) & \text{otherwise} \end{cases}$$

$\tau = 3 \cdot w(k+1) + h$ is a new counting step in the search window, a new pixel evaluated using the growing window at item k .

[0046] Finally ~~let we introduce~~ the last parameter ξ , representing the ratio between the two colors counters, is introduced once the background is identified, according to the following relationship:

$$\begin{cases} \xi = \frac{\gamma_1}{\gamma_2} & \text{if } \gamma_1 \geq \gamma_2 \Rightarrow c_1 = \text{background} \\ \xi = \frac{\gamma_2}{\gamma_1} & \text{if } \gamma_1 < \gamma_2 \Rightarrow c_2 = \text{background} \end{cases}$$

[0047] Once the algorithm is exited from the search strategy, the detection window is available to identify its content.

[0048] As mentioned above, a first condition to be satisfied, so that a region ~~would~~ can be recognized as text, is that the block has a reasonable extension. Let ~~be~~:

$$\bar{\varepsilon} = \min w$$

10 the minimum value, in terms of pixels, allowed for a region to be recognized as text region. The condition to be satisfied by a text region will be:

$$w \geq \bar{\varepsilon}$$

15

The current value fixed for the parameter $\bar{\varepsilon}$ is: $\bar{\varepsilon} = 300$.

[0049] Recalling that ξ is the ratio between the background and the text colors, a second condition to be satisfied, so that the block would be recognized as a text area, will be:

20

$$\xi \geq \bar{\xi}$$

where $\bar{\xi}$ is a modifiable parameter actually fixed as $\bar{\xi} = 1.2$. In other words:

25

if $\xi < \bar{\xi} \Rightarrow q[\cdot]$ not a text window

[0050] The block will be discarded as not a text block when one of the above conditions are not satisfied. The new search window will be $q(r, c+w)$ and it will start at position $(r, c+w)$ in the original image, or $(r+3, c)$ depending if $_$ in the previous step $_$ the end of line was reached.

[0051] Following this strategy $_$ the entire image will be scanned by the search window and text region will be detected. As text is detected $_$ the previously described post-processing operations will be applied.

[0052] Going back to Fig. 4, an input image is first subjected to a block-growing process BlGr based on whether the number of different colors does not exceed 2 ($N_c < 2$), a first indication for the presence of text. As soon as the number of colors exceeds 2, the block growing process BlGr is stopped, and the other parameters Outpar are determined, which represent the three criteria for text listed above. Based on these parameters Outpar, it is determined whether there is a text region (Txt reg ?). If so, the background color $c_{\text{background}}$ is set to white, and the text color c_{text} is set to black.

[0053] The resulting image is subjected to the scaling operation SCAL.

[0054] After the scaling operation SCAL, the text region is subjected to a thresholding operation (threshold 9), the output of

which is applied to a morphological filter (Morph. Filt.).
Thereafter, white is set back to the background color $c_{\text{background}}$, and
black is set back to the text color c_{text} . The result of this
operation forms the output image OutIm s^* that is displayed on a
5 matrix display D.

[0055] A primary aspect of the invention can be summarized as
follows. A novel technique is suggested which is able to take into
account the image content and to apply an ad-hoc scaler post-
processing only where it is requested. A viable area of application
10 of this invention is the text readability improvement in the case
of LCD devices, when, ~~and it is~~ is usually the case, we do not want
to affect other part of the displayed signal. It is, inter alia, an
object of the invention to provide an ad hoc simple text detector.
The invention proposes the design of a text detection algorithm,
15 together with a post-processing block, for text enhancement. The
invention significantly improves the performance in terms of
content readability and leads to good perceptual results of the
whole displayed signal, while keeping really low the computational
complexity of the total scaling system. The invention is preferably
20 applied in LCD scaler ICs.

[0056] It should be noted that the above-mentioned embodiments
illustrate rather than limit the invention, and that those skilled
in the art will be able to design many alternative embodiments
without departing from the scope of the appended claims. In the
25 claims, any reference signs placed between parentheses shall not be

PHIT 000001

construed as limiting the claim. The word "comprising" does not
exclude the presence of elements or steps other than those listed
in a claim. The word "a" or "an" preceding an element does not
exclude the presence of a plurality of such elements. The invention
5 can be implemented by means of hardware comprising several distinct
elements, and by means of a suitably programmed computer. In the
device claim enumerating several means, several of these means can
be embodied by one and the same item of hardware. The mere fact
that certain measures are recited in mutually different dependent
10 claims does not indicate that a combination of these measures
cannot be used to advantage.

References:

- [0057] 1.———_W.K. Pratt, "CAP 15: Morphological Image Processing" in "Digital Image Processing", J. Wiley & Sons, 1991.
- [0058] 2.———_A.V. Oppenheim, R.W. Schaffer, Discrete-Time
5 Signal Processing, Prentice-Hall International, Inc. 1989.
- [0059] 3.———_P. Zamperoni, Metodi dell'elaborazione digitale di immagini, Masson, Milano, 1990.
- [0060] 4.———_H.J.A.M. Heijmans, C. Ronse, The Algebraic Basis of Mathematical Morphology I.Dilations and Erosions, Comput.Vision
10 Graphics Image Process. 50, pp. 245-295, 1990.
- [0061] 5.———_H.J.A.M. Heijmans, C. Ronse, The Algebraic Basis of Mathematical Morphology II.Openings and Closings, CVGIP: Image Understanding, Vol. 54, No.1, pp. 74-97, 1991.
- [0062] 6.———_E.R. Dougherty, Morphological Image Processing,
15 SPIE Optical Engineering Press, 1992.
- [0063] 7.———_S.R._Sternberg, R.M._Haralick, X._Zhuang, Image Analysis Using Mathematical Morphology, IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-9, No. 4, pp. 532-550, July 1987.

PHIT 000001

ABSTRACT OF THE DISCLOSURE+

In a method of text improvement, ~~in an~~ image text in an
image is detected (SW, Det), the image is scaled (Scal) to adjust
5 first numbers of pixels per line and lines per image of the image
to second numbers of pixels per line and lines per image that fit
in with a display (D) on which the image is to be displayed, and
the image is processed (Post-proc) in dependence on a result of the
text detection.

10

~~(Fig. 4)~~